



# Neural dynamics of reward probability coding: a Magnetoencephalographic study in humans

Julie Thomas, Giovanna Vanni-Mercier and Jean-Claude Dreher\*

Cognitive Neuroscience Center, Reward and Decision-Making Team, CNRS, UMR 5229, Université de Lyon, Université Claude Bernard Lyon 1, Lyon, France

## Edited by:

Colin Camerer, California Institute of Technology, USA

## Reviewed by:

Paul Sajda, Columbia University, USA

V. S. Chandrasekhar Pammi, University of Allahabad, India

## \*Correspondence:

Jean-Claude Dreher, Cognitive Neuroscience Center, Reward and Decision-Making Team, CNRS, UMR 5229, Université de Lyon, Université Claude Bernard Lyon 1, 67 Bd Pinel, 69675 Bron, France  
e-mail: dreher@isc.cnrs.fr

Prediction of future rewards and discrepancy between actual and expected outcomes (prediction error) are crucial signals for adaptive behavior. In humans, a number of fMRI studies demonstrated that reward probability modulates these two signals in a large brain network. Yet, the spatio-temporal dynamics underlying the neural coding of reward probability remains unknown. Here, using magnetoencephalography, we investigated the neural dynamics of prediction and reward prediction error computations while subjects learned to associate cues of slot machines with monetary rewards with different probabilities. We showed that event-related magnetic fields (ERFs) arising from the visual cortex coded the expected reward value 155 ms after the cue, demonstrating that reward value signals emerge early in the visual stream. Moreover, a prediction error was reflected in ERF peaking 300 ms after the rewarded outcome and showing decreasing amplitude with higher reward probability. This prediction error signal was generated in a network including the anterior and posterior cingulate cortex. These findings pinpoint the spatio-temporal characteristics underlying reward probability coding. Together, our results provide insights into the neural dynamics underlying the ability to learn probabilistic stimuli-reward contingencies.

**Keywords:** MEG, prediction-error, reward probability coding

## INTRODUCTION

Predicting the occurrence of potentially rewarding events is a critical ability for adaptive behavior. Compelling evidence from single-unit recording in non-human primates indicate that reward probability modulates midbrain dopaminergic neurons activity both at the time of the conditioned stimuli when a prediction is made and at the time of outcome, when the discrepancy between actual and expected outcome is computed (Fiorillo et al., 2003). During appetitive classical conditioning, the phasic response of dopamine neurons increases with higher reward probability at the time of the conditioned stimulus (CS) and decreases at the time of the outcome (Fiorillo et al., 2003; Tobler et al., 2005; Kobayashi and Schultz, 2008). In humans, recent microelectrode recordings also indicate that substantia nigra neurons exhibit higher firing rates after unexpected gains than unexpected losses (Zaghloul et al., 2009). These findings support the hypotheses that midbrain dopaminergic neurons code the expected reward value at the time of the conditioned stimuli and a prediction error signal at the outcome, representing the discrepancy between anticipated and rewards effectively delivered. Building on monkey electrophysiological experiments, a number of recent fMRI studies investigated the influence of probability on reward-related brain activity (Elliott et al., 2003; Knutson et al., 2005; Abler et al., 2006; Dreher et al., 2006; Preusschoff et al., 2006; Tobler et al., 2007). Although most of these studies focused on reward probability coding in the ventral striatum, the representations of reward expectation and of reward prediction errors are not confined to subcortical regions and are also found in cingulate, prefrontal, and intra-parietal cortices (Platt

and Glimcher, 1999; Fletcher et al., 2001; Ridderinkhof et al., 2004; Sugrue et al., 2004; Dreher et al., 2006).

Much less is known concerning the representations of reward properties in early visual structures. However, recent theoretical proposals and empirical findings in humans support the view that reward properties, such as reward timing (Shuler and Bear, 2006) and prior reward history of stimuli (Serences, 2008; Gavornik et al., 2009) may be coded in areas of the visual system, including V1. Moreover, a recent fMRI study in monkeys showed that dopaminergic signals can modulate visual cortical activity, directly demonstrating that reward help to regulate selective plasticity within the visual representation of reward predicting stimuli (Arsenault et al., 2013). These studies challenge the common view that only after initial processing of low-level stimulus features in the visual cortex are higher order cortical areas engaged to process the significance of visual input and its predictive value.

The vast majority of neuroimaging studies performed in humans used fMRI to investigate the influence of reward probability on the brain. However, because of the low temporal resolution of fMRI, it has not been possible to investigate the precise timing of reward probability coding and their spatio-temporal characteristics during conditioning tasks using designs with short CS-outcome period as those used in animal experiments, nor to separate events temporally close within these tasks (Fiorillo et al., 2003; Dreher et al., 2006; Tobler et al., 2007). Other approaches using techniques with high temporal resolution, such as EEG, have focused on the neural mechanisms of feedback evaluation when subjects evaluate outcomes of their actions and use these evaluations to guide decision-making (Gehring and Willoughby,

2002; Holroyd et al., 2004; Cohen et al., 2007; Christie and Tata, 2009). However, these types of tasks may not involve the same processes as those required in classical conditioning.

Here, we report a new MEG study investigating the spatio-temporal neural dynamics of prediction and reward prediction errors computations by focusing on the timing of responses to predictive cues and rewarded/unrewarded outcomes when humans learned probabilistic associations of visual cues of slot machines with monetary rewards. We characterized the effects of reward probability on evoked magnetic fields occurring during the computations underlying reward prediction and reward prediction error.

## METHODS

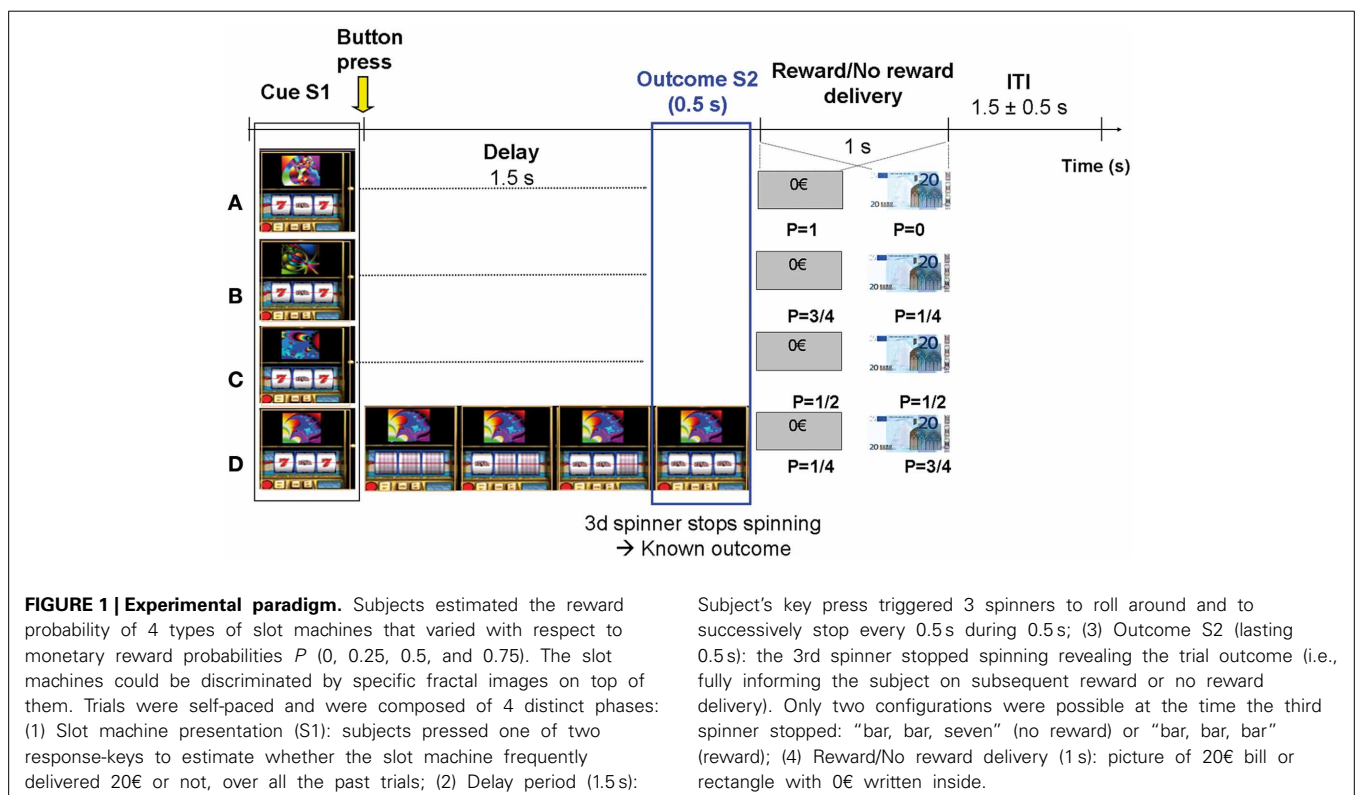
### PARTICIPANTS

Twelve right-handed subjects (seven males) participated in this MEG experiment (mean age  $\pm$  SD 23.08  $\pm$  2.23 years). They were all university students, and were free of psychiatric and neurological problems as well as of drug abuse and history of pathological gambling. Four subjects (3 males) were excluded from the data analysis because of large head movements ( $>5$  mm). All subjects gave their written, informed consent to participate in this study, which was approved by the local ethics committee. The subjects were paid for their participation, and earned extra money in proportion to their gains during the experiment.

### EXPERIMENTAL PROCEDURE

Subjects were presented with 10 runs of 4 blocks with the same elementary structure (Figure 1). In each block, one single

slot machine was presented on a computer screen during 20 consecutive trials ( $ITI = 1.5 \pm 0.5$  s). Each slot machine was made visually unique by displaying a particular fractal image on top of it. In each run, 4 types of slot machines were presented in random order and unbeknownst to the subjects were attached to 4 reward probabilities ( $P = 0; 0.25; 0.5; 0.75$ ). A total of  $10 \times 4 = 40$  different slot machines were presented. The probability of each slot machine was exact and reached at the end of each block. The paradigm is similar to the one used in a previous intra-cranial recording study (Vanni-Mercier et al., 2009). Briefly, the subjects' task was to estimate at each trial the reward probability of each slot machine at the time of its presentation, based upon all the previous outcomes of the slot machine until this trial (i.e., estimate of *cumulative* probability since the first trial). To do so, subjects had to press one of two response-buttons: one button indicating that, overall, the slot machine had a "high winning probability" and the other button indicating that overall, the slot machine had a "low winning probability." Thus, the task was *not* to predict whether the current slot machine would be rewarded or not rewarded on the current trial. Subjects were told that their current estimate had no influence on subsequent reward occurrence. During the task, subjects received no feed-back relative to their correct/incorrect estimation of the winning probability of the slot machine. Finally, at the end of each block, they were asked to classify the slot machine on a scale from 0 to 4 according to their global estimate of reward delivery over the block. The experimental paradigm was implemented with the Presentation software (<http://nbs.neuro-bs.com/presentation>).



## MEG RECORDINGS AND MRI ANATOMICAL SCANS

MEG recordings were carried out in a magnetically shielded room with a helmet shaped 275 channels gradiometer whole-head system (OMEGA; CTF Systems, VSM Medtech, Vancouver, BC, Canada). The subjects were comfortably seated upright, instructed to maintain their heads motionless and to refrain from blinking. Head motions were restricted by using a head stabilizer bladder. Between runs, subjects could have a short rest with eye blinking allowed, but were asked to stay still. Visual stimuli were projected on a translucent screen positioned 1.5 m from the subject. The subject's head position relative to MEG sensors was recorded at the beginning and end of each run using three anatomical fiducials (coils fixed at the nasion and at left and right preauricular points). These coils were also used to co-register the MEG sensor positions with the individual anatomical MRI. Subject's head position was readjusted between runs to maintain the same position all along the experiment. Subjects with head movements larger than 5 mm were excluded from the study. The mean head movement of the subjects included in the analysis was  $3.5 \pm 0.8$  mm.

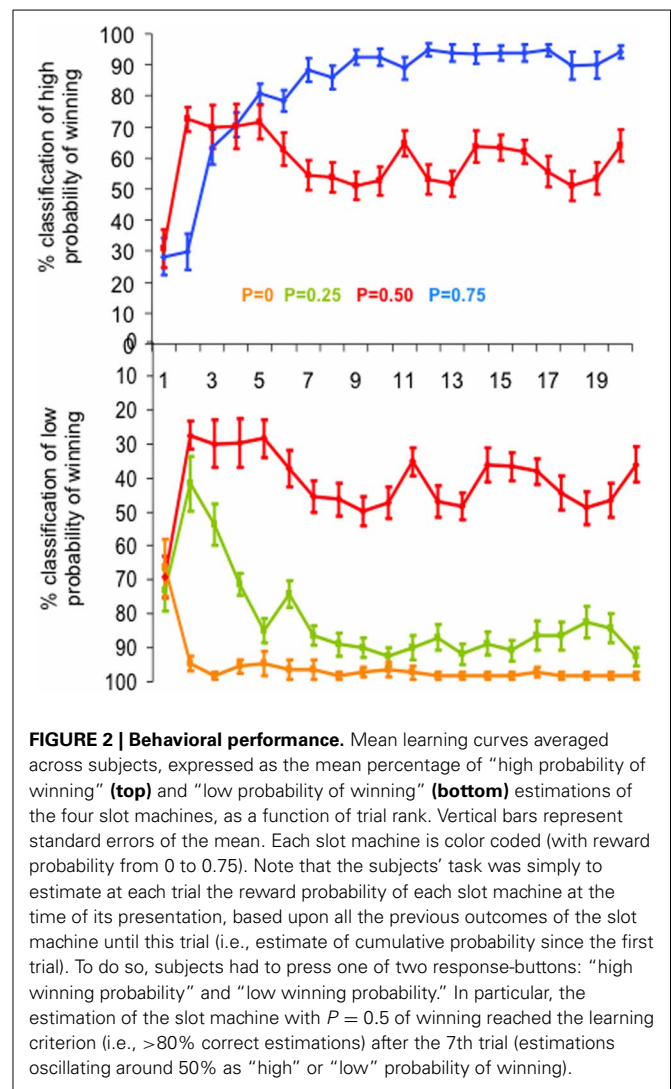
Anatomical MRI scans were obtained for each subject using a high resolution T1-weighted sequence (magnetization prepared gradient echo sequence, MP-RAGE: 176 1.0 mm sagittal slices; FOV = 256 mm, NEX = 1, TR = 1970 ms, TE = 3.93 ms; matrix =  $256 \times 256$ ; TI = 1100 ms, bandwidth = 130 Hz/pixel for 256 pixels in-plane resolution =  $1 \text{ mm}^3$ ).

MEG signals were sampled at 600 Hz with a 300 Hz cut-off filter and stored together with electro-oculogram (EOG), electrocardiogram (ECG) signals, and digital markers of specific events for subsequent off-line analysis. These markers included: 4 markers at the cue (slot machine appearance: S1) corresponding to the 4 reward probabilities of the slot machines (P0, 0.25, 0.5, and 0.75), 2 markers at the subject's behavioral responses, and 7 markers at the outcome (i.e., when the 3rd spinner stops spinning: S2), defined according to the 7 possible outcomes (3 rewarded slot machines, 3 unrewarded slot machines, and one with only unrewarded outcomes).

## BEHAVIORAL DATA ANALYSIS

We computed the percentage of correct estimations of the reward probability for each slot machine as a function of trial rank (from 1 to 20) averaged over subjects and runs (Figure 2). Estimations were defined as correct when subjects classified as "low winning" the slot machines with low reward probabilities ( $P = 0$  and  $P = 0.25$ ) and as "high winning" the slot machines with high reward probability ( $P = 0.75$ ). Since the slot machine with reward probability  $P = 0.5$  had neither "low" nor "high" winning probability and the choice was binary, the percent of 50% estimates of "high winning," or symmetrically of "low winning" probability, corresponded to the correct estimate of winning probability for this slot machine.

For the probabilities P0, P0.25, and P0.75, the trial rank when learning occurred (learning criterion) was defined as the trial rank with at least 80% of correct responses and for which the percent of correct estimation did not decrease below this limit for the remaining trials. For the probability  $P = 0.5$ , the trial rank when learning occurred was defined as the trial rank with



approximately 50% of the responses being either "high" or "low winning" probability and then oscillating around this value for the remaining trials. Moreover, results from subjects' estimation of the slot machines for each of the 20 successive presentations of a single type of slot machine within runs were compared to their classification made at the end of each block. We also analyzed the response time (RT) as a function of reward probability.

## MEG DATA PROCESSING

MEG signals were converted to a 3rd order synthetic gradient, decimated at 300 Hz, Direct Current (DC) corrected and band-pass filtered between 0.1 and 30 Hz. They were then processed with the software package for electrophysiological analyses (ELAN-Pack) developed at the INSERM U821 laboratory (Lyon, France; [http://u821.lyon.inserm.fr/index\\_en.php](http://u821.lyon.inserm.fr/index_en.php)). Trials with eye blinks or muscle activity were discarded after visual inspection, as well as cardiac artifacts, using the program DataHandler developed at the CNRS-LENA UPR 640 laboratory (Paris, France, <http://cogimage.dsi.cnrs.fr/index.htm>) (~10% of rejection in total). Signals were averaged with respect to the markers at S1 and

S2 using epochs of 3500 ms (−1500 to 2000 ms from the markers) with a 1000 ms baseline correction. Baseline activity was defined as the average activity during the ITI.

## MEG DATA ANALYSIS

### Event-related magnetic fields

For each reward probability, we ensured the statistical significance of the ERFs observed at the cue presentation (S1) and at the outcome (S2) compared to the baseline with a Wilcoxon test performed on epochs of 3500 ms, with a moving time-window of 20 ms shifted by 2 ms step.

Next, we examined the relationship between ERFs peak amplitudes and reward probability at the group level at S1 and S2, using an ANOVA with reward probability as independent factor. *Post-hoc* comparisons were then performed using Tukey's HSD tests to further assess the significant differences between ERFs peak amplitudes as a function of probability. In addition, we determined the mean onset latencies, peak latencies, and durations of the ERFs time-locked to S1 and S2.

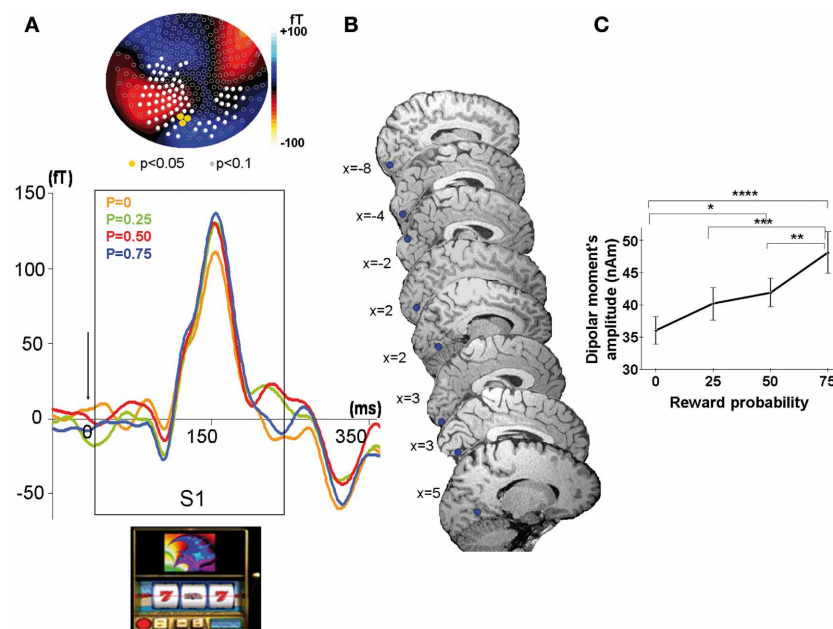
### Sources reconstruction at S1

Sources reconstruction of averaged ERFs was performed according to two different methods at S1. We used the dipolefit analysis (DipoleFit, CTF Systems, Inc., Vancouver, BC, Canada) for the ERFs observed at S1 because at the time of the cue the scalp distribution of ERFs was clearly dipolar (Figure 3A). An advantage

of this technique is that dipole pairs can be fitted to each individual dataset separately. The signal time-window used for dipole localization was 90–150 ms after S1, during the rising phase of the ERFs up to its point of maximal amplitude, because this rising phase is considered to reflect the primary source activation of the signal. No *a priori* hypothesis was made concerning the localization of dipoles necessary to explain the MEG activity recorded at the sensors level. For each source given by the default settings of the analysis software, the residual variance was calculated and the potential source was accepted if the residual variance was less than 15%. We added dipoles in a parsimonious way to reach this threshold. For each subject, 2 or 3 dipoles explained the signal with 85% of goodness-of-fit, but only the first dipole was observed in each individual subject at similar location in the visual cortex. Therefore, we considered this first dipole as the most plausible and its localization was performed in each subject using the BrainVoyager software (<http://www.brainvoyager.com>). The final criterion for the acceptance of the defined potential dipole was its physiological plausibility (location in gray matter and amplitude <250 nA m). Finally, we performed the analysis of dipole moments amplitudes as a function of reward probability at the group level with a multifactorial ANOVA.

### Sources reconstruction at S2

For the sources reconstruction of the ERFs observed at S2, we used the Synthetic Aperture Magnetometry (SAM) methodology



**FIGURE 3 | MEG response sensitive to the conditioned stimulus at the time of the cue (S1).** (A) Top: Scalp topography of the M150. Evoked related magnetic fields maps averaged across all participants showing a linear increase with reward probability at the time of the peak amplitude (=155 ms) after cue presentation (S1). Color scale indicates the intensity of the magnetic field. Color dots indicate occipital sensors (sensors MLO11, MLO12, MLO22) where the ERFs showed a significant modulation by reward probability (Yellow dots:  $P < 0.05$ ; white dots:  $P < 0.1$ ). Bottom: Grand average MEG waveform on one selected occipital sensor in the left

hemisphere (sensor MLO11) for the presentation of the slot machines (cue at S1), linearly modulated by reward probability ( $P = 0$ : yellow,  $P = 0.25$ : green,  $P = 0.5$ : red,  $P = 0.75$ : blue). (B) Source localization showing highly reproducible dipoles on sagittal slices in each subject ( $n = 8$ ). (C) Linear increase with reward probability of the dipolar moment's amplitudes of the first dipole explaining the activity elicited by the cue. Stars represent statistical differences found with a Tukey *post-hoc* between dipolar moment's amplitudes \* $p < 0.05$ ; \*\* $p < 0.005$ ; \*\*\* $p < 0.0005$ ; \*\*\*\* $p < 0.00005$ . Vertical bars represent standard errors of the mean.



(CTF Systems, Inc., Vancouver, BC, Canada) because in this case, the MEG signals were more complex (multi-peaked) and their scalp distribution was clearly multipolar. SAM is an adaptive spatial filtering algorithm or “beamformer.” The beamformer is a linear combination of the signals recorded on all sensors, optimized such that the signal estimated at a point is unattenuated while remote correlated interfering sources are suppressed. It links each voxel in the brain with the MEG sensors by constructing an optimum spatial filter for that location (Van Veen et al., 1997; Robinson and Vrba, 1999). This spatial filter is a set of weights and the source strength at the target location is computed as the weighted sum of all the MEG sensors. The algorithms of the SAM method are reported in (Hillebrand and Barnes, 2005). SAM examines the changes in signal power in a certain frequency band between two conditions for each volume element. We applied SAM to the 0.1–40 Hz frequency band to provide a statistical measure of the difference in the signal power of the two experimental conditions over the same time-window. We compared the two extreme reward probabilities at S2. That is, we focused on searching the sources showing a difference in power between  $P = 0.25$  and  $P = 0.75$  at S2 for rewarded trials (reflecting a prediction error). The time-window chosen for the analysis of the ERFs included the first ERF peak, in accordance with the significant effects observed on the sensors that is, from 0 to 300 ms. SAM analysis was applied on the whole brain volume with a 2.5 mm voxel resolution. The true  $t$ -test value was calculated at each voxel with a Jackknife approach. The Jackknife method enables accurate determination of trial-by-trial variability, while integrating the multiple covariance matrices over all but one trial. For each of these covariance matrices, SAM computes the source power on a voxel by voxel basis.

These statistical SAM data were superimposed on individual MRI data of each subject and statistical parametric maps representing significant voxels as color volumes were generated in each subject. We then performed a co-registration of the individual subjects' statistical maps on a “mean” brain obtained from all subjects. At the individual level,  $p$ -values  $< 0.05$  were considered as significant (i.e.,  $t > 2$ , Jackknife test). At the group level, the anatomical source locations were reported (Tables 1, 2) and displayed (Figure 5) if at least half of the subjects reached a statistical significance of  $t > 2$  (Jackknife test).

## RESULTS

### BEHAVIORAL RESULTS

#### Estimation of reward probability

A multifactorial ANOVA performed on the percent of correct estimates of the probability of winning (low likelihood of winning for P0 and P0.25, high likelihood of winning for P0.75 and 50% of each alternative for P0.5) showed that both reward probability ( $P$ ) and trial rank ( $R$ ) influenced the percentage of correct estimations [ $F_{P(3, 1460)} = 220.2$ ,  $p < 0.0001$ ;  $F_{R(19, 1460)} = 16.7$ ,  $p < 0.0001$ ]. The trial rank at which learning occurred depended on reward probability [interaction between trial rank and probability:  $F_{R*P(57, 1460)} = 6.9$ ,  $p < 0.001$ ]; that is, reward probability P0 reached the learning criterion (i.e.,  $>80\%$  correct estimations) after the 2nd

**Table 1 | Sources localizations of the first individual dipoles found at S1 reported in Talairach space, and their corresponding ellipsoid of confidence volume ( $p \geq 95\%$ ).**

Anatomical structures	Reward value			Volume error (cm3)
	x	y	z	
CS	−8	−87	9	0.72
CS	−4	−87	0	0.001
CS/Cuneus	−2	−86	16	0.035
CS/Cuneus	2	−87	21	0.01
CS	1	−86	−3	0.24
CS	3	−73	21	0.69
CS	3	76	19	0.59
CS	5	−65	19	0.77

Abbreviation: CS, Calcarine Sulcus.

**Table 2 | Sources localizations of ERFs found at S2 for all the subjects, in Talairach space.**

Anatomical structures	Reward prediction error		
	x	y	z
ACC	5	43	19
Post. ACC	5	−36	48
Precuneus	5	−2	54
SMG	−26	−38	44

Abbreviations: ACC, Anterior Cingulate Cortex; SMG, Supramarginal gyrus.

trial, while the reward probabilities P0.25 and 0.75 reached the learning criterion after the 5th (respectively 84.9 and 80.7% correct estimations). The reward probability P0.5 reached the learning criterion after the 7<sup>th</sup> trial (estimations oscillating around 50% as “high” or “low” probability of winning) (Figure 2).

In addition, the global estimate of reward delivery performed over each block confirmed that subjects learned the reward probability of each slot machine: they made 97.5% of correct estimations for reward probability  $P = 0$ ; 77.5% for  $P = 0.25$ ; 68.8% for  $P = 0.75$  and 82.5% for  $P = 0.5$ .

Together, these behavioral results indicate that learning of cue-outcome contingencies was performed rapidly for each type of slot machine (even for the slot machine with  $P = 0.5$ ). Thus, because the learning criterion was reached rapidly (in 2–7 trials) the effect of learning on MEG signals could not be studied and the MEG signals were analyzed after averaging all trials.

#### RESPONSE TIMES (RTs)

RTs were analyzed using a One-Way ANOVA with reward probability ( $P$ ) as independent factor. No main effect of the probability of the slot machines on RT [ $F_{(3, 8696)} = 1.4$ ,  $p = 0.238$ ] was observed. The mean RT  $\pm$  SEM for all reward probabilities and trials was:  $673.3 \pm 28.2$  ms.

## MEG SIGNALS

### MEG-evoked responses at the sensor level

**Modulation of ERFs observed at the cue (S1) by reward probability.** In each subject, strong ERFs emerged at left occipital sensors (MLO 11, 12, 22) around 90 ms after S1 (appearance of the slot machine), peaking at  $155 \text{ ms} \pm 13 \text{ ms}$  and lasting up to  $260 \text{ ms} \pm 11.4 \text{ ms}$  after the cue onset. We therefore analyzed the ERFs averaged across all subjects. For each type of slot machine (i.e., reward probability), this emerging signal was significantly different from baseline during a time window varying from 35 to 265 ms around the maximal amplitude (Wilcoxon tests,  $p$ -values varying from  $<0.0001$  to  $<0.02$  at the different sensors). At all the occipital sensors showing the ERFs at S1, there was a main effect of reward probability on the peak amplitude of these ERFs (110 to 137 fT) [ANOVA with probability as independent factor: MLO11:  $F_{(3, 5580)} = 6.3$ ,  $p < 0.0005$ ; MLO12:  $F_{(3, 5580)} = 2.8$ ,  $p < 0.05$ , MLO22:  $F_{(3, 5580)} = 2.6$ ,  $p < 0.05$ ]. Moreover, a test of linearity (Spearman) revealed that these ERFs increased linearly with reward probability at all these sensors, being minimal for P0.25 and maximal for P0.75 ( $p < 0.05$ ) (Figure 3A).

**Modulation of ERFs observed at S2 by reward probability.** Early complex ERFs emerged over occipital and temporal areas  $110 \pm 11.4 \text{ ms}$  after each successive stop of the three spinners of the slot machines, peaking  $300 \pm 16.5 \text{ ms}$  after and lasting  $450 \pm 13.2 \text{ ms}$ . Only after the third spinner stopped, giving full information about upcoming outcome (S2), was the peak amplitude of these ERFs ( $68\text{--}99 \text{ fT}$ ), observed at occipital and temporal sensors (MRO34, MRT15, MRT26, MRT27), modulated by reward probability (Figures 4A,B). This signal was significantly different from baseline for each reward probability during a time window varying from 20 to 450 ms around the maximal amplitude for rewarded trials (Wilcoxon tests,  $p$ -values varying from  $<0.0001$  to  $<0.043$ ). ANOVAs performed at the group level showed a main effect of probability. That is, reward probability modulated ERFs' peak amplitudes for rewarded trials [sensors MRO34:  $F_{\text{Probability}(2, 2173)} = 10.8$ ,  $p < 0.00005$ ; MRT 15:  $F_{\text{Probability}(2, 2173)} = 5.2$ ,  $p < 0.05$ ; MRT 26:

$F_{\text{Probability}(2, 2173)} = 7.1$ ,  $p < 0.005$ ; MRT27:  $F_{\text{Probability}(2, 2173)} = 3.6$ ,  $p < 0.05$ ]. Moreover, tests of linearity (Spearman) between ERFs' peak amplitudes and reward probability at all these sites were significant, decreasing linearly from  $P = 0.25$  to  $P = 0.75$  for rewarded trials ( $p < 0.05$ ).

## SOURCES RECONSTRUCTIONS OF ERFs

### Sources reconstruction at S1

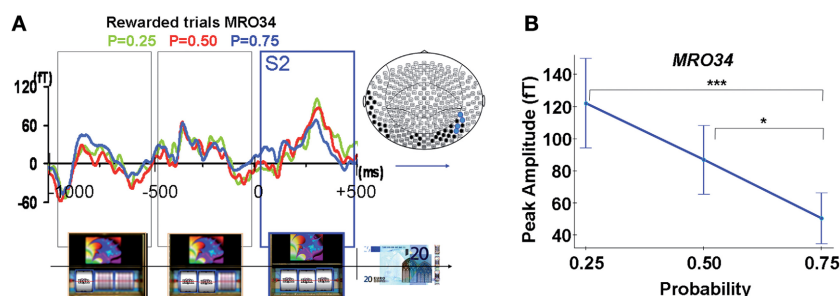
The localization of the dipole source of the ERFs observed at S1 was consistently assigned to the calcarine sulcus for 6 of the 8 subjects and to the cuneus/calcarine junction for the remaining 2 subjects (Figure 3B and Table 1). The relationship between the amplitudes of these dipolar moments averaged across subjects and reward probability monotonically increased with reward probability, being minimal for  $P = 0.25$  and maximal for  $P = 0.75$ . An ANOVA at the group level, with reward probability as independent factor showed a main effect of reward probability on dipolar moment amplitude's [ $F_{(3, 5580)} = 14.8$ ,  $p < 5.10^{-6}$ ] (Figure 3C). A subsequent test of linearity (Spearman) between dipolar moments amplitudes and reward probability revealed that these dipolar moments amplitudes increased linearly with reward probability ( $p < 0.05$ ) (Figure 3C).

### Sources reconstruction at S2

In all subjects but one, SAM activation maps identified a set of sources of the ERFs observed at the outcome S2 for rewarded trials as a function of reward probability. When comparing the sources powers between the lowest ( $P = 0.25$ ) and the highest ( $P = 0.75$ ) rewarded probability, we observed that the anterior and posterior cingulate cortices, the precuneus and the supramarginal gyrus (Brodmann's area 40) were more activated by the lowest ( $P = 0.25$ ) than the highest ( $P = 0.75$ ) rewarded probability (Figure 5 and Table 2) (Jackknife tests,  $t > 2$ ).

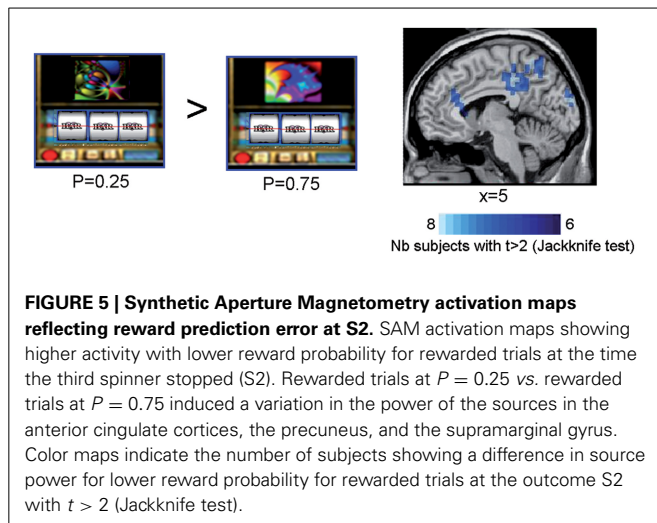
## DISCUSSION

This study used a probabilistic reward task to characterize the spatio-temporal dynamics of cerebral activity underlying reward processing using MEG. Two important results emerge from the



**FIGURE 4 | MEG response occurring during spinner's rotations and at the time of the outcome S2. (A)** Averaged signal from a representative selected occipital site (sensor MRO34), during spinners' rotation and showing a linear modulation with reward probability when the third spinner stopped (S2) for rewarded trials (Reward probability  $P = 0.25$ : green;  $P = 0.5$ : red;  $P = 0.75$ : blue). Color dots on the map's brain indicate sensors where there was a statistical difference

between peaks amplitudes for rewarded trials as a function of reward probability (blue dots:  $P < 0.05$ , black dots:  $P < 0.1$ ). **(B)** ERF's peak amplitudes from a representative selected occipital site (sensor MRO34) modulated by reward probability for unrewarded outcomes. Stars represent statistical differences (Tukey *post-hoc* between different probabilities. \* $P < 0.05$ ; \*\*\* $P < 0.0005$ ). Vertical bars represent standard errors of the mean.



present study: (1) reward value is coded early (155 ms) after the cue in the visual cortex; (2) prediction error is coded 300 ms after the outcome S2 (i.e., when the third spinner stops, fully informing subjects on subsequent reward/no reward delivery) in a network involving the anterior and posterior cingulate cortices. Moreover, reward probability modulates both ERFs coding reward prediction at the cue and prediction error at the outcome.

#### EARLY REWARD VALUE CODING AT THE TIME OF THE CUE

We found early evoked-related magnetic fields (ERFs) arising around 90 ms and peaking 155 ms after the cue presentation over the occipital cortex (M150) (**Figure 3A**). Dipole source localization showed that this early response observed at mid-occipital site was most probably generated in primary visual areas (calcarine sulcus) and to a lesser extent in secondary visual areas (**Figure 3B**). The amplitude of dipolar moments corresponding to these sources increased linearly with reward probability, indicating that visual areas code the reward value of the cue very early (**Figure 3C**). The ERFs observed at the cue are unlikely to represent the response to low level stimulus attributes of the fractal displayed on the slot machine because each of these fractals was associated with a distinct reward probability in each block, and the ERFs associated to each probability represents the mean response to all the different fractals having the same reward probability averaged over the ten runs.

Several studies have documented value-based modulations in different areas of the visual system. In rodents, primary visual cortex neurons code the timing of reward delivery (Shuler and Bear, 2006) and in humans, the visual primary cortex was found to be modulated by prior reward history in a two-choice task (Serences, 2008), showing that primary visual areas are not exclusively involved in processing low-level stimuli properties. Our results extend these findings in two important ways: first, by revealing that reward value is coded early in the primary visual cortex during a conditioning paradigm, and second, by showing that reward probability effectively modulates this early signal in a parametric way. The representation of reward value in early visual areas could be useful to increase attention to the cues having higher reward

probability and to bias decisions for high value cues. Consistent with our finding, a recent EEG/MEG study investigating reward anticipation coding reported differential activity over midline electrodes and parieto-occipital sensors (Doñamayor et al., 2012). Differences between non-reward and reward predicting cues were localized in the cuneus, also interpreted as a modulation by reward information during early visual processing.

One open question concerns whether reward probability is computed locally in the visual cortex or whether this information is received from afferent areas coding reward value even earlier than 150 ms, either directly from midbrain dopaminergic neurons or indirectly from the basal ganglia or other cortical areas. The early latency of visual cortex response to reward value (emerging at 90 ms and peaking at 150 ms) observed in the current study is compatible with direct inputs from the substantia nigra/VTa neurons that show short-latency responses (~50–100 ms) to reward delivery (Schultz, 1998) and with a recent monkey fMRI study showing that dopaminergic signals modulate visual cortex activity (Arsenault et al., 2013). In contrast, this latency would have been slower if the response resulted from polysynaptic sub-cortical or cortical routes (Thorpe and Fabre-Thorpe, 2001).

#### REWARD PREDICTION ERROR CODING AT THE OUTCOME

Although ERFs emerged over occipito-temporal sensors 90 ms after each successive stop of the three spinners of the slot machines, only after the third spinner stopped at the outcome S2—fully informing the subject on subsequent reward/no reward delivery—were these ERFs modulated by reward probability. The ERFs observed during the first two time-windows most probably reflect a motion illusion effect (**Figure 4A**), in line with the known functional specialization motion-selective units in area MT and the important role of the temporal cortex in motion perception (Kawakami et al., 2000; Senior et al., 2000; Schoenfeld et al., 2003; Sofue et al., 2003; Tanaka et al., 2007).

At the outcome, ERFs peaked around 300 ms over temporo-occipital sensors. The mean peak amplitudes of these ERFs followed the properties of a reward prediction error signal (**Figures 4B, 5**). That is, these peak amplitudes decreased linearly with reward probability for rewarded outcomes. Our results show that coding of the prediction error occurs rapidly when the third spinner stops, and before the actual reward/no reward delivery. Moreover, we disentangle the prediction error signal from the signal occurring at the time of reward delivery, often confounded in fMRI studies because of its limited temporal resolution (Behrens et al., 2008; D'Ardenne et al., 2008).

Source localization of the ERFs observed with the prediction error identified sources such as the ACC and the posterior cingulate cortex, as well as the parietal cortex. The modulation of the peak amplitude of the ERFs by reward probability observed at sites generated by ACC/posterior cingulate cortices may reflect that this signal is conveyed by dopaminergic neurons, known to exhibit discharges dependent upon reward probability (Fiorillo et al., 2003).

The ERFs observed in the current study at the outcome, peaking around 300 ms after the third spinner stopped is functionally reminiscent of both the M300, the magnetic counterpart to the

electric P300 (Kessler et al., 2005) and the magnetic equivalent of the feedback-related negative activity (mFRN). The FRN and the feedback-related P300 are two ERPs components which have been of particular interest when a feedback stimulus indicates a win or a loss to a player in a gambling game. The FRN is a fronto-central negative difference in the ERP following losses or error feedback compared to wins or positive feedback, peaking around 300 ms (Holroyd et al., 2003; Yasuda et al., 2004; Hajcak et al., 2005a,b), which may reflect a neural response to prediction errors during reinforcement learning (Holroyd and Coles, 2002; Nieuwenhuis et al., 2004). The P300, a parietally distributed ERP component, is sensitive to the processing of infrequent or unexpected events (Nieuwenhuis et al., 2005; Verleger et al., 2005; Barcelo et al., 2007). Larger P300s are elicited by negative feedback when participants thought they made a correct response, by positive feedback when participants thought they made an incorrect response (Horst et al., 1980). Our results are in accordance with previous reports of reduced amplitude of either the FRN (Holroyd et al., 2009, 2011; Potts et al., 2011; Doñamayor et al., 2012) or the ensuing P300 (Hajcak et al., 2005a,b) following expected compared to unexpected non-rewarding outcomes.

As the prediction error signal found in the current study, both the FRN and P300 evoked potentials are sensitive to stimulus probability (Nieuwenhuis et al., 2005; Verleger et al., 2005; Barcelo et al., 2007; Cohen et al., 2007; Mars et al., 2008). However, our conditioning paradigm markedly differs from decision making paradigms classically used to observe feedback-related signals (Falkenstein et al., 1991; Nieuwenhuis et al., 2004; Cohen et al., 2007; Bellebaum and Daum, 2008; Christie and Tata, 2009; Cavanagh et al., 2010), making systematic functional comparisons difficult. For the same reason, it is difficult to compare our results with those of a recent MEG study using a risk aversion paradigm in which subjects had to choose between two risky gambles leading to potential losses and gains (Talmi et al., 2012). The analysis of this study, limited to the sensor level, reported a signal that resembled the FRN emerging around 320 ms after outcome. Although this study argued that the presence of a condition with losses may strengthen conclusions regarding reward prediction error, potential losses can also induce counterfactual effects (Mellers et al., 1997), which is a problem when the same reference point (i.e., no gain) is not included in each gamble (Breiter et al., 2001). Indeed, the emotional response to the outcome of a gamble depends not only on the obtained outcome but also on its alternatives (Mellers et al., 1997). Thus, the neural mechanisms of feedback evaluation after risky gambles likely involve different processes as those engaged in a simple classical conditioning paradigm. Yet, all these paradigms indicate to the subjects the difference between prediction and actual outcomes and participate to different forms of reinforcement learning and evaluation of outcomes of decisions to guide reward-seeking behavior.

Further studies are needed to dissociate the different roles of feedback in risky decision making, in classical and instrumental conditioning and in social decision making, involving both common and specific neural mechanisms. For example, concerning the counterfactual effect mentioned above, recent fMRI and MEG studies investigated brain responses involved in the feeling

of regret and disappointment by manipulating the feedback participants saw after making a decision to play certain gambles: full-feedback (regret) vs. partial-feedback (disappointment: when only the outcome from chosen gamble is presented) (Coricelli and Rustichini, 2010; Giorgetta et al., 2013). Another question relates to whether the FRN and ACC activities express salience prediction errors rather than reward prediction errors, as suggested by recent EEG and fMRI data using different types of rewards and punishments (Metereau and Dreher, 2013; Talmi et al., 2013). Studies manipulating different rewards and punishments are needed to clarify this question (but see Dreher, 2013; Sescousse et al., 2013).

## CONCLUSION

The current study identified the spatio-temporal characteristics underlying reward probability coding in the human brain. It provides evidence that the brain computes separate signals, at a sub-second time scale, in successive brain areas along a temporal sequence when expecting a potential reward. First, processing of reward value at the cue takes place at an early stage in visual cortical areas, and then the anterior and posterior cingulate cortices, together with the parietal cortex compute a prediction error signal sensitive to reward probability at the time of the outcome. Together, these results suggest that these signals are necessary when expecting a reward and when learning probabilistic stimuli-outcome associations. Our findings provide important insights into the neurobiological mechanisms underlying the ability to code reward probability, to predict upcoming rewards and to detect changes between these predictions and rewards effectively delivered.

## ACKNOWLEDGMENTS

This work was performed within the framework of the LABEX ANR-11-LABX-0042 of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). This work was funded by grants from the “Fondation pour la Recherche Médicale,” the DGA and the Fyssen Foundation to Jean-Claude Dreher. Many thanks to the MEG CERMEP staff (C. Delpuech and F. Lecaigard), and to Dr O. Bertrand and P-E. Aguera for helpful assistance.

## REFERENCES

- Abler, B., Walter, H., Erk, S., Kammerer, H., and Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage* 31, 790–795. doi: 10.1016/j.neuroimage.2006.01.001
- Arsenault, J. T., Nelissen, K., Jarraya, B., and Vanduffel, W. (2013). Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron* 77, 1174–1186. doi: 10.1016/j.neuron.2013.01.008
- Barcelo, F., Perianez, J. A., and Nyhus, E. (2007). An information theoretical approach to task-switching: evidence from cognitive brain potentials in humans. *Front. Hum. Neurosci.* 1:13. doi: 10.3389/neuro.09.013.2007
- Behrens, T. E., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. (2008). Associative learning of social value. *Nature* 456, 245–249. doi: 10.1038/nature07538
- Bellebaum, C., and Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *Eur. J. Neurosci.* 27, 1823–1835. doi: 10.1111/j.1460-9568.2008.06138.x
- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639. doi: 10.1016/S0896-6273(01)00303-8



- Cavanagh, J. F., Frank, M. J., Klein, T. J., and Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* 49, 3198–3209. doi: 10.1016/j.neuroimage.2009.11.080
- Christie, G. J., and Tata, M. S. (2009). Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage* 48, 415–422. doi: 10.1016/j.neuroimage.2009.06.076
- Cohen, M. X., Elger, C. E., and Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* 35, 968–978. doi: 10.1016/j.neuroimage.2006.11.056
- Coricelli, G., and Rustichini, A. (2010). Counterfactual thinking and emotions: regret and envy learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 241–247. doi: 10.1098/rstb.2009.0159
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., and Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319, 1264–1267. doi: 10.1126/science.1150605
- Doñamayor, N., Schoenfeld, M. A., and Münte, T. F. (2012). Magneto- and electroencephalographic manifestations of reward anticipation and delivery. *Neuroimage* 62, 17–29. doi: 10.1016/j.neuroimage.2012.04.038
- Dreher, J. C., Kohn, P., and Berman, K. F. (2006). Neural coding of distinct statistical properties of reward information in humans. *Cereb. Cortex* 16, 561–573. doi: 10.1093/cercor/bhj004
- Dreher, J. C. (2013). Neural coding of computational factors affecting decision making. *Prog. Brain Res.* 202, 289–320. doi: 10.1016/B978-0-444-62604-2.00016-2
- Elliott, R., Newman, J. L., Longe, O. A., and Deakin, J. F. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J. Neurosci.* 23, 303–307.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455. doi: 10.1016/0013-4694(91)90062-9
- Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902. doi: 10.1126/science.1077349
- Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., et al. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat. Neurosci.* 4, 1043–1048. doi: 10.1038/nn733
- Gavornik, J. P., Shuler, M. G., Loewenstein, Y., Bear, M. F., and Shouval, H. Z. (2009). Learning reward timing in cortex through reward dependent expression of synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6826–6831. doi: 10.1073/pnas.0901835106
- Gehring, W. J., and Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282. doi: 10.1126/science.1066893
- Giorgetta, C., Greucci, A., Bonini, N., Coricelli, G., Demarchi, G., Braun, C., et al. (2013). Waves of regret: a MEG study of emotion and decision-making. *Neuropsychologia* 51, 38–51. doi: 10.1016/j.neuropsychologia.2012.10.015
- Hajcak, G., Moser, J. S., Yeung, N., and Simons, R. F. (2005a). On the ERN and the significance of errors. *Psychophysiology* 42, 151–160. doi: 10.1111/j.1469-8986.2005.00270.x
- Hajcak, G., Holroyd, C. B., Moser, J. S., and Simons, R. F. (2005b). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology* 42, 161–170. doi: 10.1111/j.1469-8986.2005.00278.x
- Hillebrand, A., and Barnes, G. R. (2005). Beamformer analysis of MEG data. *Int. Rev. Neurobiol.* 68, 149–171. doi: 10.1016/S0074-7742(05)68006-3
- Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709. doi: 10.1037/0033-295X.109.4.679
- Holroyd, C. B., Larsen, J. T., and Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology* 41, 245–253. doi: 10.1111/j.1469-8986.2004.00152.x
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., and Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport* 14, 2481–2484. doi: 10.1097/00001756-200312190-00037
- Holroyd, C. B., Krigolson, O. E., and Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport* 22, 249–252. doi: 10.1097/WNR.0b013e328345441d
- Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., and Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cogn. Affect. Behav. Neurosci.* 9, 59–70. doi: 10.3758/CABN.9.1.59
- Horst, R. L., Johnson, R. Jr., and Donchin, E. (1980). Event-related brain potentials and subjective probability in a learning task. *Mem. Cognit.* 8, 476–488. doi: 10.3758/BF03211144
- Kawakami, O., Kaneoke, Y., and Kakigi, R. (2000). Perception of apparent motion is related to the neural activity in the human extrastriate cortex as measured by magnetoencephalography. *Neurosci. Lett.* 285, 135–138. doi: 10.1016/S0304-3940(00)01050-8
- Kessler, K., Schmitz, F., Gross, J., Hommel, B., Shapiro, K., and Schnitzler, A. (2005). Target consolidation under high temporal processing demands as revealed by MEG. *Neuroimage* 26, 1030–1041. doi: 10.1016/j.neuroimage.2005.02.020
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812. doi: 10.1523/JNEUROSCI.0642-05.2005
- Kobayashi, S., and Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* 28, 7837–7846. doi: 10.1523/JNEUROSCI.1600-08.2008
- Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., et al. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* 28, 12539–12545. doi: 10.1523/JNEUROSCI.2925-08.2008
- Mellers, B. A., Schwartz, A., Ho, K., and Ritov, I. (1997). Decision affect theory: emotional reactions to the outcomes of risky options. *Psychol. Sci.* 8, 423–429. doi: 10.1111/j.1467-9280.1997.tb00455.x
- Metereau, E., and Dreher, J. C. (2013). Cerebral correlates of salient prediction error for different rewards and punishments. *Cereb. Cortex* 23, 477–87. doi: 10.1093/cercor/bhs037
- Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532. doi: 10.1037/0033-2909.131.4.510
- Nieuwenhuis, S., Holroyd, C. B., Mol, N., and Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neurosci. Biobehav. Rev.* 28, 441–448. doi: 10.1016/j.neubiorev.2004.05.003
- Platt, M. L., and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238. doi: 10.1038/22268
- Potts, G. F., Martin, L. E., Kamp, S.-M., and Donchin, E. (2011). Neural response to action and reward prediction errors: comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations. *Psychophysiology* 48, 218–228. doi: 10.1111/j.1469-8986.2010.01049.x
- Preusschoff, K., Bossaerts, P., and Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390. doi: 10.1016/j.neuron.2006.06.024
- Ridderinkhof, K. R., van den Wildenberg, W. P., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016
- Robinson, S. E., and Vrba, J. (1999). “Functional neuroimaging by synthetic aperture magnetometry,” in *Recent Advances in Biomagnetism*, eds T. Yoshimoto, M. Kotani, S. Kuriki, H. Karibe, and N. Nakasato (Sendai: Tohoku University Press), 302–305.
- Schoenfeld, M. A., Woldorff, M., Duzel, E., Scheich, H., Heinze, H. J., and Mangun, G. R. (2003). Form-from-motion: MEG evidence for time course and processing sequence. *J. Cogn. Neurosci.* 15, 157–172. doi: 10.1162/089892903321208105
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Curr. Biol.* 10, 16–22. doi: 10.1016/S0960-9822(99)00259-6
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron* 60, 1169–1181. doi: 10.1016/j.neuron.2008.10.051

- Sescousse, G., Caldú, X., Segura, B., and Dreher, J. C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 681–96. doi: 10.1016/j.neubiorev.2013.02.002
- Shuler, M. G., and Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609. doi: 10.1126/science.1123513
- Sofue, A., Kaneoke, Y., and Kakigi, R. (2003). Physiological evidence of interaction of first- and second-order motion processes in the human visual system: a magnetoencephalographic study. *Hum. Brain Mapp.* 20, 158–167. doi: 10.1002/hbm.10138
- Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787. doi: 10.1126/science.1094765
- Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., and Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *Neuroimage* 59, 635–645. doi: 10.1016/j.neuroimage.2011.06.051
- Talmi, D., Atkinson, R., and El-Dereby, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *J. Neurosci.* 33, 8264–8269. doi: 10.1523/JNEUROSCI.5695-12.2013
- Tanaka, E., Noguchi, Y., Kakigi, R., and Kaneoke, Y. (2007). Human cortical response to various apparent motions: a magnetoencephalographic study. *Neurosci. Res.* 59, 172–182. doi: 10.1016/j.neures.2007.06.1471
- Thorpe, S. J., and Fabre-Thorpe, M. (2001). Neuroscience. Seeking categories in the brain. *Science* 291, 260–263. doi: 10.1126/science.1058249
- Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645. doi: 10.1126/science.1105370
- Tobler, P. N., O'Doherty, J. P., Dolan, R. J., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632. doi: 10.1152/jn.00745.2006
- Van Veen, B. D., van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880. doi: 10.1109/10.623056
- Vanni-Mercier, G., Mauguier, F., Isnard, J., and Dreher, J. C. (2009). The hippocampus codes the uncertainty of cue-outcome associations: an intracranial electrophysiological study in humans. *J. Neurosci.* 29, 5287–5294. doi: 10.1523/JNEUROSCI.5298-08.2009
- Verleger, R., Gorgen, S., and Jaskowski, P. (2005). An ERP indicator of processing relevant gestalts in masked priming. *Psychophysiology* 42, 677–690. doi: 10.1111/j.1469-8986.2005.354.x
- Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., and Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport* 15, 2561–2565. doi: 10.1097/00001756-200411150-00027
- Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., et al. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science* 323, 1496–1499. doi: 10.1126/science.1167342

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 July 2013; paper pending published: 17 September 2013; accepted: 28 October 2013; published online: 18 November 2013.

Citation: Thomas J, Vanni-Mercier G and Dreher J-C (2013) Neural dynamics of reward probability coding: a Magnetoencephalographic study in humans. *Front. Neurosci.* 7:214. doi: 10.3389/fnins.2013.00214

This article was submitted to *Decision Neuroscience*, a section of the journal *Frontiers in Neuroscience*.

Copyright © 2013 Thomas, Vanni-Mercier and Dreher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.